Interindividual Variation in Nucleotide Excision Repair Genes and Risk of Endometrial Cancer

Jocelyn M. Weiss,[1,2,4] Noel S. Weiss,[1,2] Cornelia M. Ulrich,[1,2] Jennifer A. Doherty,[1] Lynda F. Voigt,[1,2] and Chu Chen[1,2,3]

[1]Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center; [2]Department of Epidemiology, School of Public Health and Community Medicine; [3]Department of Otolaryngology: Head and Neck Surgery, School of Medicine, University of Washington, Seattle, Washington; and [4]Currently at the Division of Cancer Epidemiology and Genetics, NCI/NIH, Bethesda, Maryland

Abstract

Exposure to estrogens is a likely cause of endometrial cancer, but the means by which estrogens exert this effect are not entirely clear. One hypothesis is that certain estrogen metabolites bind to the DNA, forming bulky adducts that damage the DNA and initiate carcinogenesis. A woman's reduced capacity to repair such damage may increase her risk of endometrial cancer. We conducted a population-based case-control study in western Washington State to address the role of variation in nucleotide excision repair genes on the risk of endometrial cancer. Case women (n = 371), ages 50 to 69 years, were diagnosed with invasive endometrial cancer between 1994 and 1999. Control women (n = 420) were selected using random-digit dialing (ages 50-65 years) and by random selection from Health Care Financing Administration data files (ages 66-69 years). Genotyping assays were done for *ERCC1*, *ERCC2* (*XPD*), *ERCC4* (*XPF*), *ERCC5* (*XPG*), *XPA*, and *XPC*. No appreciable differences between cases and controls were observed in the genotype distributions of *ERCC1* (c8092a and c19007t), *ERCC2* (D312N, K751Q, and c22541a), *ERCC4* (R415Q and t30028c), or *ERCC5* (D1104H). Carriage of at least one variant allele for *XPA G23A* was associated with decreased risk of endometrial cancer [odds ratio (OR), 0.70; 95% confidence interval (95% CI), 0.53-0.93]. Carriage of at least one *XPC* A499V variant allele was associated with a modest decrease in risk (OR, 0.79; 95% CI, 0.59-1.05). Women with variant alleles at both *XPC* A499V and K939Q had 58% of the risk of women with no *XPC* variant alleles (OR, 0.58; 95% CI, 0.35-0.96). Our data suggest that interindividual variation in *XPA* and *XPC* influences a woman's risk of endometrial cancer. (Cancer Epidemiol Biomarkers Prev 2005;14(11):2524–30)

Introduction

Although endogenous and exogenous estrogens are unequivocally associated with the risk of endometrial cancer (reviewed in refs. 1-3), the mechanisms by which estrogens might act to influence the development of this disease remain unclear. One widely accepted hypothesis emphasizes the role of unopposed estrogen in stimulating cellular proliferation (4, 5). A complementary hypothesis proposes that estrogens and their metabolites play a role in tumor initiation via direct damage to the DNA by the formation of bulky DNA adducts (6).

The major estrogens in women, estradiol and estrone, undergo oxidative metabolism through hydroxylation at various sites, but the major pathways are 2-hydroxylation, 4-hydroxylation, and 16α-hydroxylation. The 2-OH and 4-OH estrogens can be further oxidized to semiquinones and quinones, which can form bulky DNA adducts (7) and can undergo redox cycling, producing reactive oxygen species that may cause oxidative stress, lipid peroxidation (8), and DNA damage (6, 9-11). The catechol estrogen quinones can be deactivated by conjugation with glutathione (12) or converted to hydroquinones, which protect against redox cycling (13, 14).

Interindividual variation in genes that govern the structure and/or activity of the enzymes involved in estrogen biosynthesis, catabolism, and response are thought to play a role in a woman's susceptibility to endometrial cancer. With respect to DNA damage response, the role of genetic variation in the mismatch repair pathway on risk of hereditary nonpolyposis colorectal cancer–associated endometrial cancer has been extensively studied, but germ line mutations in mismatch repair enzymes (MSH2, MLH1, MLH6, and PMS2) have rarely been observed in sporadic endometrial cancers (15-17). It is hypothesized that the DNA damage caused either by oxidative stress, and/or by estrogen metabolites binding to the DNA to form bulky DNA adducts, along with a woman's inability to repair this damage, is a possible mechanism by which estrogens increase risk.

The nucleotide excision repair (NER) pathway, which is comprised of >30 gene products, removes bulky chemical adducts, UV-induced pyrimidine dimers, and cross-links (18-20). In addition to estrogen, other common environmental and dietary exposures, such as cigarette smoke and well-cooked meat, contribute to bulky adduct formation in the cell (21). NER involves the recognition of DNA distortion caused by the presence of a bulky adduct by the RAD23B-XPC (xeroderma pigmentosum complementation group C) heterodimer followed by binding of XPA/RPA to the lesion. XPA is thought to be crucial for the subsequent positioning of the ERCC1-ERCC4 (XPF) endonuclease (22). Helicases ERCC3 (XPB) and ERCC2 (XPD; two of the subunits of transcription factor IIH) unwind the DNA helix, and the ERCC5 (XPG) and ERCC1-ERCC4 nucleases excise a 24- to 32-bp segment containing the bulky adduct at its 3′ and 5′ ends, respectively. The resultant gap is filled by DNA polymerase (POLE or POLD1) and ligase I (LIG1) using the intact strand as a template (23). It is not entirely clear whether *in vivo* repair

Received 6/5/05; revised 8/19/05; accepted 9/9/05.

Grant support: NIH grants R35-CA39779 (N.S. Weiss), R01-CA75977 (N.S. Weiss), R03-CA80636 (C. Chen), R01-CA105212 (C. Chen), N01-HD23166 (Women's Contraceptive and Reproductive Experiences), and K05-CA92002 (N.S. Weiss) and Fred Hutchinson Cancer Research Center.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Cancer Epidemiology Biomarkers and Prevention (http://cebp.aacrjournals.org/).

Requests for reprints: Jocelyn M. Weiss, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Occupational and Environmental Epidemiology Branch, 6120 Executive Boulevard, EPS 8123, MSC 7240, Rockville, MD 20852. Phone: 301-435-2358; Fax: 301-402-1819. E-mail: weissjoc@mail.nih.gov

doi:10.1158/1055-9965.EPI-05-0414

involves sequential loading assembly of individual factors or loading of a complete repair complex onto a DNA lesion, but in either case, the repair factors are likely to act in a defined order (18-20, 22).

We sought to test the hypothesis that interindividual variation in the genes involved in the repair of estrogen DNA adducts is associated with risk of endometrial cancer. Using a candidate gene approach, we selected 11 common (≥5% minor allele frequency) single nucleotide polymorphisms (SNP) in six NER genes (*ERCC1, ERCC2, ERCC4, ERCC5, XPA*, and *XPC*), including all amino acid changes known at that time, and evaluated the associations between the respective genotypes and risk of endometrial cancer. We additionally investigated whether certain combinations of genotypes are associated with a particularly altered risk of disease.

Materials and Methods

Study Population. This study, described in Doherty et al. (24), is comprised of women who took part in a population-based case-control study of endometrial cancer, as well as a subset of controls from a population-based case-control study of breast cancer (the multisite Women's Contraceptive and Reproductive Experiences [CARE] Study; ref. 25). Case women included White and Black female residents of King, Pierce, and Snohomish counties of western Washington state ages 50 to 69 years diagnosed with invasive endometrial cancer between January 1, 1994 and December 31, 1995 (King County only) and between July 1, 1997 and December 31, 1999. These women were identified through the Cancer Surveillance System, a population-based tumor registry affiliated with the Surveillance, Epidemiology, and End Results Program of the National Cancer Institute (26). Cases ages 50 to 65 years were included only if they had a residential telephone in their homes, and those ages 66 to 69 years were included only if they were present in the Health Care Financing Administration data files used to select controls. Interviews were obtained from 472 (81.1%) of eligible cases, with blood samples provided by 383 (65.8%) of the eligible women (81.1% of interviewed women).

Control women with intact uteri and no prior history of endometrial cancer were selected from the three-county area using random-digit dialing (ages 50-65 years; ref. 27) and also by random selection from yearly Health Care Financing Administration data files (ages 66-69 years). The controls were frequency matched to the cases by 5-year age group and county of residence. The overall random-digit dialing response (the screening response multiplied by the interview response) was 76.3%, with 297 of eligible random-digit dialing controls interviewed. Of the 175 eligible Health Care Financing Administration controls, 116 (66.3%) agreed to an interview.

Eligible control women from the King County site of the CARE breast cancer study (25), conducted during the same period of time, were included in the endometrial cancer study. The overall levels of screening and interview response for King County were 83.6% and 88.3%, respectively. Of the 132 King County CARE control women ages 50 to 64 years, with intact uteri, who were invited to provide a blood sample, we successfully obtained a blood sample from 115 (87.1%). Overall, of the 930 eligible controls, 665 (71.5%) were interviewed and 450 (48.4%) provided a blood sample (67.7% of interviewed controls).

The data from one control in the earlier (CARE) study, who was subsequently ascertained as a case in the later ENRICH study, was included in both the case and control groups; one case was excluded due to poor quality interview data; and the DNA samples of four controls could not be analyzed. Additionally, there were only 11 cases and 26 controls who were Hispanic or non-Caucasian. To minimize the potential confounding influence of race/ethnicity, our current analysis is restricted to non-Hispanic Caucasian women, for a total of 371 cases and 420 controls.

In-person interviews were conducted according to a standard protocol after informed consent. Participants were asked only about events that occurred before their ''reference date.'' For cases, this was the month and year of their endometrial cancer diagnosis. Reference date for the CARE controls was the month and year of their screening interview. All other controls were assigned reference years based on the distribution of diagnosis years among the cases and assigned a random reference month. A calendar of life events was constructed to facilitate recall. Data collected include demographic factors; height; weight at different ages; reproductive, contraceptive, and menstrual history; family history of cancer; history of selected chronic conditions; and history of contraceptive and noncontraceptive hormone use. Color pictures of oral contraceptive and hormone replacement therapy pill packs were used to prompt a detailed history of use.

Laboratory Assays. Genomic DNA was extracted from peripheral blood samples using a salt precipitation method (28). Genotyping assays were done by multiplex single-base-pair primer extension reactions with fluorescent-labeled dideoxynucleotide triphosphates (SNaPshot) on an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) and by RFLP analysis. Laboratory personnel were blinded to the case-control status. Quality control (positive and negative) samples were included in each batch of samples to validate genotyping procedures. Concordance for blind replicates (4% of total samples) was 100%. Scoring of the genotyping results was done independently by two molecular biologists, with any discrepancies resolved by retesting of the samples.

Gene-specific PCR products containing the SNP for each candidate gene of interest were amplified using the PCR primers and cycling conditions listed in Supplementary Table 1.[5]

Determination of Genotypes by SNaPshot Chemistry (Applied Biosystems). After amplification, the PCR products to undergo SNaPshot (*ERCC1 C8092A, ERCC1 C19007T, ERCC2* D312N, *ERCC4* R415Q, *ERCC4 T30028C, ERCC5* D1104H, *XPA G23A, XPC* A499V, and *XPC T30028C*) were treated with 5 units shrimp alkaline phosphatase (U.S. Biochemical Corp., Cleveland, OH) and 2 units exonuclease I (New England Biolabs, Ipswich, MA) at 37°C for 1 hour followed by heat inactivation at 75°C for 15 minutes to avoid participation of primers and unincorporated nucleotides in the subsequent SNaPshot primer-extension reaction. Each SNaPshot reaction contained 3 μL of PCR products, 5 μL of 2× SNaPshot Multiplex Ready Reaction Mix, and 0.02 to 0.1 μmol/L of each probe. Cycling of the SNaPshot reaction was as follows: 25 cycles of 96°C for 10 seconds, 48°C for 5 seconds, and 60°C for 30 seconds. After cycling, this reaction was treated with 1 unit shrimp alkaline phosphatase for 1 hour at 37°C followed by heat inactivation at 75°C for 15 minutes; 0.5 μL of this multiplex reaction was combined with 0.1 μL of LIZ-120 size standard (Applied Biosystems) and 13.4 μL of formamide, denatured, and subjected to capillary electrophoresis on an ABI 3100 Genetic Analyzer. Genotype determinations were made using GeneScan and GeneMapper (Applied Biosystems) software.

Determination of SNP Genotypes by RFLP. Following amplification, the PCR products to undergo RFLP (*ERCC2 C22541A* and *ERCC2* K751Q) were digested with restriction enzymes known to cut differentially based on presence of

[5] Supplementary data available at http://cebp.aacrjournals.org.

the wild-type or variant alleles. Amplified *ERCC2 C22541A* products were digested with *Tfi*I, which cut the wild-type allele at 587 and 57 bp and the variant allele at 474, 113, and 57 bp. Amplified *ERCC2* K751Q products were digested with *Pst*I, which cut the wild-type allele at 161 and 41 bp and the variant allele at 120 and 41 bp.

Statistical Analyses

Logistic regression analysis was used to compute odds ratios (OR) and associated 95% confidence intervals (95% CI) relating each of the SNPs as well as combinations of SNPs to the risk of endometrial cancer. When examining SNP combinations, biologically relevant exposure variables were created by gene (*ERCC1, ERCC2*, and *ERCC4*), heterodimer complex (*ERCC1-ERCC4*), and assembly of the DNA repair machinery (*ERCC1-ERCC4-XPA, XPA-ERCC5, XPA-ERCC2*, and *ERCC5-ERCC1-ERCC4*). Only matching variables and factors that altered the ORs by ≥10% were included in the final multivariate models. Among controls, the comparability of the distribution of genotypes for a given SNP with a prediction based on Hardy-Weinberg equilibrium was assessed using the Pearson χ^2 test with one degree of freedom. All analyses were done using the STATA statistical package (version 8; STATA Corp., College Station, TX). Haplotype analyses were done using the HPlus software package (version 2.1; Fred Hutchinson Cancer Research Center, Seattle, WA; refs. 29, 30), which employs Expectation-Maximization with a modified progressive ligation computational algorithm to infer haplotypes. Haplotype analyses allowed us to combine information from multiple, individually genotyped, markers that have been inherited together on a single chromosome. If a genetic characteristic that is truly associated with disease risk is present in the region captured by the genotyped markers, then the haplotypes should capture the variation directly at that locus. As this was a study with no available family data, it was necessary to infer haplotypes for individuals in the study by use of partially observed phase information using a computer-based algorithm (29, 30). The HPlus program handles data with missing genotypes, allows for adjustment of relevant covariates (matching variables, environmental factors, lifestyle variables, etc.), and allows for evaluation of effect modification when assessing haplotype-disease associations. Regardless of the inference method used, it is important to note that the absence of phase information can lead to uncertainty in haplotype assignment, which in turn can lead to biased estimates of effect (31). Diplotype (haplotype pair) analyses were done using the PHASE estimation software (version 2; University of Washington, Seattle, WA). The results, inferred diplotypes (haplotype pairs) and their probabilities for each individual, were imported into STATA. Clustered logistic regression, using the probabilities assigned to each diplotype as weights in the regression, was used to compute ORs relating diplotypes to risk of endometrial cancer. This method of approximating joint likelihood for haplotype risk variables and haplotype frequencies has been described elsewhere for matched case-control data (31).

Results

The characteristics of the study population (Table 1) have been described previously (24). The mean age for both cases and controls was 59 years. Relative to controls, women with endometrial cancer were more likely to be obese, have fewer than two children, have a history of diabetes or hypertension, never have used oral contraceptives, to have taken postmenopausal hormone therapy in ways known to increase risk, and to not have smoked cigarettes.

Table 1. Characteristics of the study population

	Cases (n = 371), n (%)	Controls (n = 420), n (%)
Age (y)		
50-54	94 (25.3)	130 (31.0)
55-59	103 (27.8)	103 (24.5)
60-64	97 (26.1)	101 (24.0)
65-69	77 (20.8)	86 (20.5)
Education		
Less than high school	35 (9.4)	29 (6.9)
High school grad	92 (24.8)	84 (20.0)
Some college/tech school	127 (34.2)	133 (31.7)
More than or graduate school	117 (31.5)	174 (41.4)
No. births		
0	65 (17.5)	43 (10.2)
1	46 (12.4)	35 (8.3)
≥2	260 (70.1)	342 (81.4)
Body mass index (kg/m^2)		
<18.5	9 (2.4)	7 (1.7)
18.5-24.9	126 (34.3)	232 (55.4)
25.0-29.9	99 (27.0)	109 (26.0)
30.0-34.9	53 (14.4)	45 (10.7)
≥35	80 (21.8)	26 (6.2)
Missing	4 (1.1)	1 (0.2)
History of diabetes*		
No	327 (88.0)	398 (95.0)
Yes	44 (12.0)	21 (5.0)
Missing	0 (0.0)	1 (0.2)
History of hypertension*		
No	240 (64.7)	309 (73.7)
Yes	131 (35.3)	110 (26.2)
Missing	0 (0.0)	1 (0.2)
Use of oral contraceptives		
Never	171 (46.3)	170 (40.5)
Ever	198 (53.6)	236 (56.2)
Missing	2 (0.5)	14 (3.3)
Use of postmenopausal estrogens† (mo)		
0-6	165 (73.0)	190 (91.3)
≥6	61 (27.0)	18 (8.7)
Missing	145 (39.1)	212 (50.5)
Cigarette smoking ‡		
Never	201 (60.7)	179 (51.6)
Former	130 (39.3)	168 (48.4)
Current	40 (10.8)	73 (17.4)

*History of diabetes or hypertension as reported by participant, regardless of medication use.
†Postmenopausal estrogen use represents use for at least 6 months of unopposed estrogens or estrogens opposed by progestogens for <10 d/mo; women with at least 6 month of use of other forms of postmenopausal hormone therapy are excluded from this analysis.
‡Current cigarette smoking defined as that within the 1 year before reference date.

For the SNPs under study, allele frequencies were consistent with those reported for other Caucasian populations. All observed genotype frequencies among controls were statistically consistent with the Hardy-Weinberg equilibrium. The genotype distributions in cases and controls of the 11 studied NER polymorphisms are shown in Table 2. For the *XPA G23A* polymorphism, the presence of at least one variant allele was associated with a modest decrease in risk of endometrial cancer (OR, 0.70; 95% CI, 0.53-0.93). The presence of at least one variant allele for the *XPC* A499V was associated with a borderline decreased risk of endometrial cancer (OR, 0.79; 95% CI, 0.59-1.05). The observed ORs for both *XPA G23A* and *XPC* A499V were suggestive of a trend between increasing number of variant alleles and the magnitude of decreased risk. No appreciable differences between cases and controls were observed in the genotype distributions of *ERCC1 C8092A, ERCC1 C19007T, ERCC2* D312N, *ERCC2* K751Q, *ERCC2 C22541A, ERCC4* R415Q, *ERCC4 T30028C*, or *ERCC5* D1104H. The presence of at least one variant allele at each of the two *XPC* loci was associated with a decreased risk of

endometrial cancer (OR, 0.58; 95% CI, 0.35-0.96; Table 3). Adjustment for potential confounders, including postmenopausal hormone use and body mass index, the two strongest risk factors for endometrial cancer, did not appreciably alter any of the observed estimates of association.

The NER machinery is thought to assemble in a sequential manner, with many of the proteins physically interacting with one another. ERCC1 and ERCC4 form a heterodimeric complex that is responsible for the 5′ incision of DNA at the site of damage. XPA interacts with many of the core repair proteins and is essential for the correct positioning of the ERCC1-ERCC4 endonuclease (32-35). No association was observed between genotype for the *ERCC1-ERCC4* heterodimeric complex and risk of endometrial cancer (data not shown). However, when evaluating the combined *ERCC1-ERCC4* and *XPA* genotypes, reductions in risk were observed among those individuals with at least one *XPA G23A* variant allele and at least one variant allele at zero (OR, 0.77; 95% CI, 0.36-1.65), two (OR, 0.52; 95% CI, 0.30-0.90), or three (OR, 0.35; 95% CI, 0.16-0.80) sites in *ERCC1-ERCC4* but not those with a variant allele at one (OR, 1.25; 95% CI, 0.65-2.40) or four sites (OR, 0.98; 95% CI, 0.22-4.36; Table 4). Variation in other examined gene combinations (*ERCC1-ERCC4-XPA, XPA-ERCC5, XPA-ERCC2*, and *ERCC5-ERCC1-ERCC4*) was not associated with risk (data not shown).

The least common *XPC* haplotype among cases and controls (variant allele for *XPC* A499V and no variant allele for *XPC* K939Q) was associated with a marginally decreased risk of endometrial cancer (OR, 0.82; 95% CI, 0.64-1.05; Table 5). Additionally, compared with women with the wild-type *XPC* diplotype (499AA and 939KK), women with all other diplotype combinations had a marginally decreased risk of

Table 2. Distribution of nucleotide excision repair genotypes among cases and controls

Gene/polymorphism	No. cases (%), n = 371	No. controls (%), n = 420	OR (95% CI)*
ERCC1 C8092A			
CC	205 (55.3)	235 (56.0)	Reference
CA/AA	166 (44.7)	185 (44.0)	0.99 (0.75-1.32)
CA	142 (38.3)	159 (37.9)	1.00 (0.74-1.34)
AA	24 (6.5)	26 (6.2)	0.97 (0.54-1.75)
ERCC1 C19007T			
CC	134 (36.1)	170 (40.5)	Reference
CT/TT	237 (63.9)	250 (59.5)	1.16 (0.86-1.55)
CT	188 (50.7)	196 (46.7)	1.19 (0.88-1.61)
TT	49 (13.2)	54 (12.9)	1.04 (0.66-1.64)
ERCC2 D312N (*G* > *A*)			
DD	152 (41.0)	186 (44.3)	Reference
DN/NN	219 (59.0)	234 (55.7)	1.13 (0.85-1.51)
DN	173 (46.6)	176 (41.9)	1.19 (0.88-1.61)
NN	46 (12.4)	58 (13.8)	0.96 (0.61-1.50)
ERCC2 K751Q (*A* > *C*)			
KK	142 (38.3)	159 (37.9)	Reference
KQ/QQ	229 (61.7)	261 (62.1)	0.96 (0.72-1.28)
KQ	181 (48.8)	197 (46.9)	1.01 (0.74-1.37)
QQ	48 (12.9)	64 (15.2)	0.81 (0.52-1.25)
ERCC2 C22541A			
CC	117 (31.5)	137 (32.6)	Reference
CA/AA	254 (68.4)	283 (67.4)	1.06 (0.78-1.43)
CA	188 (50.7)	207 (49.3)	1.06 (0.77-1.46)
AA	66 (17.8)	76 (18.1)	1.03 (0.68-1.56)
ERCC4 R415Q (*G* > *A*)			
RR	316 (85.2)	369 (87.9)	Reference
RQ/QQ	55 (14.8)	51 (12.1)	1.23 (0.82-1.86)
RQ	54 (14.6)	49 (11.7)	1.27 (0.83-1.92)
QQ	1 (0.3)	2 (0.5)	0.48 (0.04-5.43)
ERCC4 T30028C			
TT	186 (50.1)	214 (51.0)	Reference
TC/CC	185 (49.9)	206 (49.0)	1.01 (0.76-1.34)
TC	160 (43.1)	171 (40.7)	1.05 (0.78-1.41)
CC	25 (6.7)	35 (8.3)	0.84 (0.48-1.47)
ERCC5 D1104H (*G* > *C*)			
DD	215 (58.0)	250 (59.5)	Reference
DH/HH	156 (42.0)	170 (40.5)	1.07 (0.80-1.42)
DH	134 (36.1)	148 (35.2)	1.07 (0.79-1.43)
HH	22 (5.9)	22 (5.2)	1.06 (0.56-2.00)
XPA G23A			
GG	195 (52.6)	185 (44.0)	Reference
GA/AA	176 (47.4)	235 (56.0)	0.70 (0.53-0.93)
GA	147 (39.6)	191 (45.5)	0.72 (0.54-0.97)
AA	29 (7.8)	44 (10.5)	0.62 (0.37-1.04)
XPC A499V (*C* > *T*)			
AA	211 (56.9)	213 (50.7)	Reference
AV/VV	160 (43.1)	207 (49.3)	0.79 (0.59-1.05)
AV	129 (34.8)	166 (39.5)	0.80 (0.59-1.08)
VV	31 (8.4)	41 (9.8)	0.76 (0.46-1.28)
XPC K939Q (*A* > *C*)			
KK	153 (41.2)	164 (39.0)	Reference
KQ/QQ	218 (58.8)	256 (61.0)	0.89 (0.67-1.19)
KQ	153 (41.2)	198 (47.1)	0.81 (0.60-1.11)
QQ	65 (17.5)	58 (13.8)	1.17 (0.76-1.78)

*Adjusted for age at reference date (50-54, 55-59, 60-64, and 65-69 y) and county of residence (King, Pierce, and Snohomish).

Table 3. Risk of endometrial cancer in relation to the number of rare alleles within a given gene

Gene*	Frequency		OR (95% CI)†
	Cases (%), n = 371	Controls (%), n = 420	
ERCC1 (*C8092A* and *C19007T*)			
0	131 (35.3)	165 (39.3)	Reference
1	77 (20.8)	75 (17.9)	1.27 (0.86-1.89)
2	163 (43.9)	180 (42.9)	1.09 (0.79-1.49)
ERCC2 (D312N, K751Q, and *C22541A*)			
0-1	128 (34.5)	154 (36.7)	Reference
2	150 (40.4)	160 (38.1)	1.10 (0.79-1.52)
3	93 (25.1)	106 (25.2)	1.03 (0.72-1.49)
ERCC4 (R415Q and *T30028C*)			
0	186 (50.1)	214 (51.0)	Reference
1	130 (35.0)	155 (36.9)	0.95 (0.70-1.29)
2	55 (14.8)	51 (12.1)	1.21 (0.78-1.86)
XPC (A499V and K939Q)			
0	59 (15.9)	48 (11.4)	Reference
1	246 (66.3)	281 (66.9)	0.70 (0.46-1.07)
2	66 (17.8)	91 (21.7)	0.58 (0.35-0.96)

*Values represent the number of sites with at least one rare allele across each of the respective SNPs for each gene.
†Adjusted for age at reference date (50-54, 55-59, 60-64, and 65-69 y) and county of residence (King, Pierce, and Snohomish).

Table 5. Association between NER haplotypes and endometrial cancer, by gene

Haplotypes*	Frequency		OR (95% CI)†
	Cases	Controls	
ERCC1 (*C8092A* and *C19007T*)			
00	0.60	0.63	Reference
10	0.01	0.01	1.48 (0.49-4.46)
01	0.14	0.12	1.19 (0.87-1.61)
11	0.24	0.24	1.01 (0.79-1.29)
ERCC2 (D312N, K751Q, and *C22541A*)			
000	0.15	0.16	Reference
100	0.06	0.04	1.48 (0.88-2.50)
010	0.06	0.06	0.99 (0.59-1.67)
001	0.41	0.40	1.06 (0.79-1.43)
110	0.29	0.30	1.00 (0.73-1.36)
011	0.02	0.02	0.96 (0.44-2.10)
ERCC4 (R415Q and *T30028C*)			
00	0.72	0.71	Reference
01	0.21	0.22	0.92 (0.72-1.18)
11	0.08	0.06	1.16 (0.77-1.75)
XPC (A499V and K939Q)			
00	0.36	0.33	Reference
10	0.26	0.29	0.82 (0.64-1.05)
01	0.38	0.37	0.93 (0.74-1.17)

*For each haplotype, 0 = common allele and 1 = rare allele for each SNP (in the order listed).
†Adjusted for age at reference date (50-54, 55-59, 60-64, and 65-69 y) and county of residence (King, Pierce, and Snohomish).

endometrial cancer (OR, 0.69; 95% CI, 0.46-1.04; Table 6). Carriage of at least one VK haplotype (variant allele for *XPC* A499V and no variant allele for *XPC* K939Q) was associated with a 38% reduction in risk when compared with women with the *XPC* AK/AK diplotype and a 20% reduction in risk when compared with women with all other diplotypes (OR, 0.62; 95% CI, 0.40-0.96 and OR, 0.79; 95% CI, 0.59-1.04, respectively; data not shown). Carriage of an infrequent *ERCC2* haplotype (variant allele for *ERCC2* D312N and wild-type for *ERCC2* K751Q and *ERCC2 C22541A*) was associated with a marginally increased risk (OR, 1.48; 95% CI, 0.88-2.50; data not shown). No differences in haplotype distributions were observed between cases and controls for the *ERCC1* and *ERCC4* genes. No differences in haplotype distribution between cases and controls were observed for chromosome 19, which contains both the *ERCC1* and *ERCC2* genes.

Discussion

This is the first study to examine the role of genetic variation in nucleotide excision repair genes and risk of endometrial cancer. A study by Han et al. (36) observed no association between two genes in the double-strand break repair pathway, *XRCC2* (R188H) or *XRCC3* (5′ untranslated region *4541 A → G*, IVS5-14 *17893 A → G*, T241M), and risk of endometrial cancer.

In this study, the presence of the *XPA G23A* variant allele was associated with a modest decrease in the risk of endometrial cancer. Previous studies have examined the *XPA G23A* genotype only in relation to risk of lung cancer. Park et al. (37) and Wu et al. (38) reported 40% to 60% increased risks of lung cancer among Koreans and Caucasian Americans with carriage of the variant allele. In a study by Butkiewicz et al. (39), neither the heterozygous nor homozygous variant genotype was associated with risk of lung adenocarcinoma when compared with the homozygous wild type. In a recent study, the *XPA 23AA* genotype was associated with a borderline increased risk of lung cancer (aa versus gg: OR, 1.7; 95% CI, 1.0-3.1), but the *XPA 23GG* genotype was associated with a borderline increased risk when combined with the *XPC* 939KK genotype (*GG*/QQ versus *AA*/KK: OR, 3.1; 95% CI, 1.0-9.9; ref. 40). With respect to studies that evaluated DNA repair phenotypes in relation to *XPA* genotype, the data are limited. Wu et al. (38) used the host cell reactivation assay and observed significantly higher mean DNA repair capacity among controls with at least one copy of the *XPA G23A* wild-type allele (9.53 ± 4.81) when compared with those with no copies (8.29 ± 2.81, $P = 0.03$). A corresponding difference in repair capacity was not observed among case patients.

Table 4. Risk of endometrial cancer, by *ERCC1-ERCC4* heterodimer and *XPA* genotypes

ERCC1-ERCC4*	XPA *G23A*	No. cases (%), n = 371	No. controls (%), n = 420	OR (95% CI)†
0	*GG*	36 (57.1)	39 (45.9)	Reference
	GA/AA	27 (42.9)	46 (54.1)	0.77 (0.36-1.65)
1	*GG*	40 (46.5)	52 (50.5)	Reference
	GA/AA	46 (53.5)	51 (49.5)	1.25 (0.65-2.40)
2	*GG*	72 (54.1)	55 (41.0)	Reference
	GA/AA	61 (45.9)	79 (59.0)	0.52 (0.30-0.90)
3	*GG*	35 (53.8)	25 (35.2)	Reference
	GA/AA	30 (46.2)	46 (64.8)	0.35 (0.16-0.80)
4	*GG*	12 (50.0)	14 (51.8)	Reference
	GA/AA	12 (50.0)	13 (48.2)	0.98 (0.22-4.36)

**ERCC1-ERCC4* heterodimer: values represent the number of sites with at least one rare allele across the two genes of the heterodimeric complex (*ERCC1 8092, ERCC1 19007, ERCC4* 415, and *ERCC4 30028*).
†Adjusted for age at reference date (50-54, 55-59, 60-64, and 65-69 y) and county of residence (King, Pierce, and Snohomish).

Table 6. Risk of endometrial cancer associated with *XPC* diplotype (haplotype pair)

Haplotype 1		Haplotype 2		Frequency		OR (95% CI)*
A499V	K939Q	A499V	K939Q	Cases	Controls	
A	K	A	K	0.16	0.12	1.00 (Reference)
All other diplotypes combined †				0.84	0.88	0.69 (0.46-1.04)
A	K	A	Q	0.24	0.25	0.65 (0.40-1.04)
A	K	V	K	0.17	0.18	0.68 (0.41-1.13)
A	K	V	Q	0.11	0.02	3.79 (0.44-32.72)
A	Q	A	Q	0.17	0.14	0.89 (0.52-1.50)
A	Q	V	Q	0.27	0.22	0.89 (0.06-13.54)
A	Q	V	K	0.17	0.21	0.58 (0.35-0.97)
V	K	V	K	0.08	0.10	0.62 (0.34-1.14)

*Adjusted for age at reference date (50-54, 55-59, 60-64, and 65-69 y) and county of residence (King, Pierce, and Snohomish).
†All possible diplotypes other than AK/AK.

We observed slight reductions in the risk associated with the presence of the variant allele at *XPC* A499V and *XPC* K939Q, and a larger reduction when the respective variant alleles were present at both loci. A trend of decreasing ORs among women with zero, one, or two variant alleles in these two *XPC* loci was observed, suggesting the possibility of incomplete or codominant inheritance. The reduction in risk associated with carriage of *XPC* variant alleles continued to be evident when haplotype and diplotype analyses were considered. Three prior epidemiologic studies of the A499V polymorphism have not observed an association with the risk of melanoma (41) or lung cancer (42, 43). However, another study reported an increased risk of lung cancer among individuals with the variant allele (AV/VV versus AA: OR, 1.6; 95% CI, 1.1-2.2; ref. 44). The K939Q polymorphism has been associated with an increased risk of bladder (QQ versus KK: OR, 2.0; 95% CI, 1.1-3.6; ref. 45) and lung cancer (QQ versus KK/KQ: OR, 1.8; 95% CI, 1.1-3.1; ref. 46). In another study of lung cancer, the presence of at least one rare allele was associated with a decreased risk of small cell carcinoma (KQ/QQ versus KK: OR, 0.6; 95% CI, 0.4-1.0) but not of other histologic subtypes (42). Vogel et al. (40) recently reported a 3-fold increased risk of lung cancer associated with carriage of the combined *XPC* 939QQ but only when combined with the *XPA* 23 homozygous common or homozygous variant genotype (40). No association was observed between *XPC* 939 genotype and lung cancer in another study, except when combined with *XPC* A249V genotype (43). Forsti et al. (47) observed a borderline decreased risk of breast cancer with at least one XPC 939Q allele in a Polish population (OR, 0.8; 95% CI, 0.6-1.0) and no association in a Finnish population, whereas Zhang et al. (48) observed a borderline increased risk with carriage of the heterozygous genotype (OR, 1.5; 95% CI, 1.0-2.2). No association was observed for *XPC* 939 genotype in a study of basal cell carcinoma (49).

The *XPC* K939Q polymorphism is reported to be in linkage disequilibrium with a polyintronic $(AT)_n$ (PAT+/−) polymorphism (50), which has been associated with lower DNA repair capacity (as measured by the host cell reactivation assay) and significantly increased risks of squamous cell carcinoma of the head and neck (51). However, a significant difference in the frequency of chromosomal abnormalities by K939Q genotype was not observed in two Czech studies of healthy volunteers and tire plant workers (52, 53). Functional evidence has yet to be presented for the *XPC* A499V polymorphism.

Our study has some noteworthy limitations. First, its relatively modest sample size allowed us to reliably identify only moderate-to-strong associations. Second, for a given gene, we studied no more than three SNPs. The absence of an association between our chosen SNPs and risk of disease does not preclude a potential role of a given gene in the NER pathway, or the pathway as a whole, and the incidence of endometrial cancer. The particular SNPs examined in this study were chosen using a candidate gene approach, with our intent to include SNPs from multiple genes in the NER pathway and to include known common (≥5%) SNPs with amino acid changes. We are aware of the limitations of this selection method, particularly in light of the more recent developments in tagSNP and haplotype approaches. As they are generally more powerful for detecting associations than using single markers, these approaches tend to be preferred for characterizing interindividual variation across genes. Ideally, we would have chosen SNPs known to be functional. However, particularly at the time of SNP selection, there were only a limited number of studies examining the phenotypic effect of variation in the NER pathway.

There are several possible bases for our observation of an association between endometrial cancer risk and *XPA G23A* and *XPC* genotypes. First, the findings could represent true inverse associations between carriage of *XPA* and *XPC* variant alleles and risk of endometrial cancer. Second, in light of the multiple comparisons being made, we cannot distinguish between the possibility of a chance association and a genuine one. Third, the alleles observed to be associated with a decreased risk of endometrial cancer might be in linkage disequilibrium with other allele(s) that have a true etiologic effect.

This study suggests avenues for further research. The *XPA* and *XPC* genotypes observed here to be inversely associated with risk of endometrial cancer have previously been reported to be positively associated with the risk of lung and bladder cancer but possibly with a reduced risk of breast cancer. Evaluation in other studies of the possible association between *XPA* and *XPC* polymorphisms and all these forms of cancer, including elucidation of the functional effect of variation in these genes, is very much in order.

References

1. Cook LS, Weiss NS. Endometrial cancer. In: Goldman MB, Hatch MC, editors. Women and health. London: Academic Press; 1999. p. 916–31.
2. Grady D, Ernster VL. Endometrial cancer. In: Schottenfeld D, Fraumeni JF, editors. Cancer epidemiology and prevention. New York (NY): Oxford University Press; 1996. p. 1058–89.
3. Voigt LF, Weiss NS. Epidemiology of endometrial cancer. Cancer Treat Res 1989;49:1–21.
4. Key TJ, Pike MC. The dose-effect relationship between 'unopposed' oestrogens and endometrial mitotic rate: its central role in explaining and predicting endometrial cancer risk. Br J Cancer 1988;57:205–12.
5. Siiteri PK. Steroid hormones and endometrial cancer. Cancer Res 1978; 38:4360–6.
6. Zhu BT, Conney AH. Functional role of estrogen metabolism in target cells: review and perspectives. Carcinogenesis 1998;19:1–27.
7. Cavalieri EL, Stack DE, Devanesan PD, et al. Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators. Proc Natl Acad Sci U S A 1997;94:10937–42.
8. Wang MY, Liehr JG. Induction by estrogens of lipid peroxidation and lipid peroxide-derived malonaldehyde-DNA adducts in male Syrian hamsters: role of lipid peroxidation in estrogen-induced kidney carcinogenesis. Carcinogenesis 1995;16:1941–5.
9. Cavalieri E, Frenkel K, Liehr JG, Rogan E, Roy D. Estrogens as endogenous genotoxic agents: DNA adducts and mutations. J Natl Cancer Inst Monogr 2000;27:75–93.
10. Yager JD. Endogenous estrogens as carcinogens through metabolic activation. J Natl Cancer Inst Monogr 2000;27:67–73.
11. Yager JD, Liehr JG. Molecular mechanisms of estrogen carcinogenesis. Annu Rev Pharmacol Toxicol 1996;36:203–32.
12. Butterworth M, Lau SS, Monks TJ. Formation of catechol estrogen glutathione conjugates and γ-glutamyl transpeptidase-dependent nephrotoxicity of 17β-estradiol in the golden Syrian hamster. Carcinogenesis 1997; 18:561–7.
13. Lind C, Hochstein P, Ernster L. DT-diaphorase as a quinone reductase: a cellular control device against semiquinone and superoxide radical formation. Arch Biochem Biophys 1982;216:178–85.
14. Schulz WA, Krummeck A, Rosinger I, et al. Increased frequency of a null-allele for NAD(P)H: quinone oxidoreductase in patients with urological malignancies. Pharmacogenetics 1997;7:235–9.
15. Kobayashi K, Matsushima M, Koi S, et al. Mutational analysis of mismatch repair genes, hMLH1 and hMSH2, in sporadic endometrial carcinomas with microsatellite instability. Jpn J Cancer Res 1996;87:141–5.

16. Lim PC, Tester D, Cliby W, et al. Absence of mutations in DNA mismatch repair genes in sporadic endometrial tumors with microsatellite instability. Clin Cancer Res 1996;2:1907–11.
17. Banno K, Susumu N, Yanokura M, et al. Association of HNPCC and endometrial cancers. Int J Clin Oncol 2004;9:262–9.
18. de Laat WL, Jaspers NG, Hoeijmakers JH. Molecular mechanism of nucleotide excision repair. Genes Dev 1999;13:768–85.
19. Araujo SJ, Wood RD. Protein complexes in nucleotide excision repair. Mutat Res 1999;435:23–33.
20. Costa RM, Chigancas V, Galhardo Rda S, Carvalho H, Menck CF. The eukaryotic nucleotide excision repair pathway. Biochimie 2003;85:1083–99.
21. Wogan GN, Hecht SS, Felton JS, Conney AH, Loeb LA. Environmental and chemical carcinogenesis. Semin Cancer Biol 2004;14:473–86.
22. Volker M, Mone MJ, Karmakar P, et al. Sequential assembly of the nucleotide excision repair factors *in vivo*. Mol Cell 2001;8:213–24.
23. Lindahl T, Wood RD. Quality control by DNA repair. Science 1999;286: 1897–905.
24. Doherty JA, Weiss NS, Freeman RJ, et al. Genetic factors in catechol estrogen metabolism in relation to the risk of endometrial cancer. Cancer Epidemiol Biomarkers Prev 2005;14:357–66.
25. Marchbanks PA, McDonald JA, Wilson HG, et al. The NICHD Women's Contraceptive and Reproductive Experiences Study: methods and operational results. Ann Epidemiol 2002;12:213–21.
26. Hankey BF, Ries LA, Edwards BK. The Surveillance, Epidemiology, and End Results program: a national resource. Cancer Epidemiol Biomarkers Prev 1999;8:1117–21.
27. Waksberg J. Sampling methods for random digit dialing. J Am Statistical Soc 1978;73:40–6.
28. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988; 16:1215.
29. Li SS, Khalid N, Carlson C, Zhao LP. Estimating haplotype frequencies and standard errors for multiple single nucleotide polymorphisms. Biostatistics 2003;4:513–22.
30. Zhao LP, Li SS, Khalid N. A method for the assessment of disease associations with single-nucleotide polymorphism haplotypes and environmental variables in case-control studies. Am J Hum Genet 2003;72:1231–50.
31. Kraft P, Cox DG, Paynter RA, Hunter D, De Vivo I. Accounting for haplotype uncertainty in matched association studies: a comparison of simple and flexible techniques. Genet Epidemiol 2005;28:261–72.
32. Mone MJ, Bernas T, Dinant C, et al. *In vivo* dynamics of chromatin-associated complex formation in mammalian nucleotide excision repair. Proc Natl Acad Sci U S A 2004;101:15933–7.
33. Saijo M, Matsuda T, Kuraoka I, Tanaka K. Inhibition of nucleotide excision repair by anti-XPA monoclonal antibodies which interfere with binding to RPA, ERCC1, and TFIIH. Biochem Biophys Res Commun 2004;321: 815–22.
34. Rademakers S, Volker M, Hoogstraten D, et al. Xeroderma pigmentosum group A protein loads as a separate factor onto DNA lesions. Mol Cell Biol 2003;23:5755–67.
35. de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NG, Hoeijmakers JH. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. Genes Dev 1998;12: 2598–609.
36. Han J, Hankinson SE, Hunter DJ, De Vivo I. Genetic variations in XRCC2 and XRCC3 are not associated with endometrial cancer risk. Cancer Epidemiol Biomarkers Prev 2004;13:330–1.
37. Park JY, Park SH, Choi JE, et al. Polymorphisms of the DNA repair gene xeroderma pigmentosum group a and risk of primary lung cancer. Cancer Epidemiol Biomarkers Prev 2002;11:993–7.
38. Wu X, Zhao H, Wei Q, et al. XPA polymorphism associated with reduced lung cancer risk and a modulating effect on nucleotide excision repair capacity. Carcinogenesis 2003;24:505–9.
39. Butkiewicz D, Popanda O, Risch A, et al. Association between the risk for lung adenocarcinoma and a (−4) G-to-A polymorphism in the XPA gene. Cancer Epidemiol Biomarkers Prev 2004;13:2242–6.
40. Vogel U, Overvad K, Wallin H, Tjonneland A, Nexo BA, Raaschou-Nielsen O. Combinations of polymorphisms in XPD, XPC and XPA in relation to risk of lung cancer. Cancer Lett 2005;222:67–74.
41. Blankenburg S, Konig IR, Moessner R, et al. No association between three xeroderma pigmentosum group C and one group G gene polymorphisms and risk of cutaneous melanoma. Eur J Hum Genet 2005;13:253–5.
42. Lee GY, Jang JS, Lee SY, et al. XPC polymorphisms and lung cancer risk. Int J Cancer 2005;115:807–13.
43. Shen M, Berndt SI, Rothman N, et al. Polymorphisms in the DNA nucleotide excision repair genes and lung cancer risk in Xuan Wei, China. Int J Cancer 2005;116:768–73.
44. Hu Z, Wang Y, Wang X, et al. DNA repair gene XPC genotypes/haplotypes and risk of lung cancer in a Chinese population. Int J Cancer 2005;115: 478–83.
45. Sanyal S, Festa F, Sakano S, et al. Polymorphisms in DNA repair and metabolic genes in bladder cancer. Carcinogenesis 2004;25:729–34.
46. Blankenburg S, Konig IR, Moessner R, et al. Assessment of 3 xeroderma pigmentosum group C gene polymorphisms and risk of cutaneous melanoma: a case-control study. Carcinogenesis 2005 [Epub ahead of print].
47. Forsti A, Angelini S, Festa F, et al. Single nucleotide polymorphisms in breast cancer. Oncol Rep 2004;11:917–22.
48. Zhang L, Zhang Z, Yan W. Single nucleotide polymorphisms for DNA repair genes in breast cancer patients. Clin Chim Acta 2005;359:150–5.
49. Festa F, Kumar R, Sanyal S, et al. Basal cell carcinoma and variants in genes coding for immune response, DNA repair, folate and iron metabolism. Mutat Res 2005;574:105–11.
50. Khan SG, Metter EJ, Tarone RE, et al. A new xeroderma pigmentosum group C poly(AT) insertion/deletion polymorphism. Carcinogenesis 2000;21: 1821–5.
51. Shen H, Sturgis EM, Khan SG, et al. An intronic poly (AT) polymorphism of the DNA repair gene XPC and risk of squamous cell carcinoma of the head and neck: a case-control study. Cancer Res 2001;61:3321–5.
52. Vodicka P, Kumar R, Stetina R, et al. Markers of individual susceptibility and DNA repair rate in workers exposed to xenobiotics in a tire plant. Environ Mol Mutagen 2004;44:283–92.
53. Vodicka P, Kumar R, Stetina R, et al. Genetic polymorphisms in DNA repair genes and possible links with DNA repair rates, chromosomal aberrations and singlestrand breaks in DNA. Carcinogenesis 2004;25:757–63.